U.S. DEPARTMENT OF COMMERCE
National Technical Information Service

PB-265 313

Optimization of State Water Quality Monitoring Systems

Corvallis Environmental Research Lab, Oreg

1976

ERL-Cor-007J

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO. EPA-600/J-76-029	2. JOURNAL ARTICLE
3.	PB 265 313
4. TITLE AND SUBTITLE Optimization of State Water Quality Monitoring Systems	5. REPORT DATE August 1976
	6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) Donald H. Lewis	8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Assessment & Criteria Development Division, Ecosystems Modeling & Analysis Branch, Corvallis Environmental Research Laboratory, Office of Research & Development, U.S. EPA, Corvallis, Oregon 97330	10. PROGRAM ELEMENT NO. 1HA091
	11. CONTRACT/GRANT NO.
12. SPONSORING AGENCY NAME AND ADDRESS Same as above.	13. TYPE OF REPORT AND PERIOD COVERED Journal article
	14. SPONSORING AGENCY CODE EPA/600/02

15. SUPPLEMENTARY NOTES CONTACT: D.H. Lewis, 8-420-4754 (Corvallis, OR)
Published in Journal of Computers and Operations Research, Vol. 3, No. 2-3, pp. 127-143, Pergamon Press, 1976, August

16. ABSTRACT

The water quality monitoring activities required of the States under the Federal Water Pollution Control Act Ammendments (PL92-500) will require resources well beyond what is likely to be available, if current systems structures and operating procedures are followed. This paper describes some quantitative procedures developed to optimize monitoring resource allocation, with the goal of getting the greatest possible return on a limited monitoring budget.

The ambient and effluent standards which form the basis of the monitoring system are described. Stochastic inputs to the river basin which lead to violations of these standards are identified, and progress in characterizing these variations mathematically is discussed.

Quantitative priority-setting procedures for the two principal monitoring activities (ambient and effluent), which take into account these stochastic inputs, are reviewed. Both procedures are designed to tell the state agency planner where the next increment of resource should be expended to maximize alternative monitoring objective functions.

Approaches to integrating these two primary monitoring functions and opportunities for further applications of operations research techniques to improve program effectiveness and efficiency are suggested.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Water Pollution Control Act monitoring systems PL92-500 water quality	detecting violation of water quality standards monitoring resource allocation	06/F, 08/H

18. DISTRIBUTION STATEMENT Release Unlimited	19. SECURITY CLASS (This Report) unclassified	21. NO. OF PAGES 19
	20. SECURITY CLASS (This page) unclassified	22. PRICE PC A02 MF A01

EPA Form 2220-1 (9-73)

REPRODUCED BY
NATIONAL TECHNICAL
INFORMATION SERVICE
U.S. DEPARTMENT OF COMMERCE
SPRINGFIELD, VA. 22161

Journal of Computers and Operations
Research, Vol. 3, No. 2-3, 127-143,
August 1976

EPA-600/J-76-029
JOURNAL ARTICLE

Comput. & Ops Res., Vol. 3, pp. 127-143. Pergamon Press, 1976. Printed in Great Britain

OPTIMIZATION OF STATE WATER QUALITY MONITORING SYSTEMS

DONALD H. LEWIS*

Office of Research and Development, U.S. Environmental Protection Agency, Corvallis, OR 97330, U.S.A.

Scope and purpose—The purpose of this article is to describe the development and use of optimizatioin procedures which indicate where and how often to sample water quality to maximize the chance of detecting violation of water quality standards. Two types of sampling network are discussed: (1) those intended to detect instream violation, and (2) those oriented toward point-of-discharge violation.

Abstract—The water quality monitoring activities required of the States under the Federal Water Pollution Control Act Ammendments (PL92-500) will require resources well beyond what is likely to be available, if current systems structures and operating procedures are followed. This paper describes some quantitative procedures developed to optimize monitoring resource allocation, with the goal of getting the greatest possible return on a limited monitoring budget.

The ambient and effluent standards which form the basis of the monitoring system are described. Stochastic inputs to the river basin which lead to violations of these standards are identified, and progress in characterizing these variations mathematically is discussed.

Quantitative priority-setting procedures for the two principal monitoring activities (ambient and effluent), which take into account these stochastic inputs, are reviewed. Both procedures are designed to tell the state agency planner where the next increment of resource should be expended to maximize alternative monitoring objective functions.

Approaches to integrating these two primary monitoring functions and opportunities for further applications of operations research techniques to improve program effectiveness and efficiency are suggested.

1. INTRODUCTION

The Federal Water Pollution Control Act Amendments (PL92-500) establishes as a national goal achievement of fishable/swimmable surface waters throughout the nation by 1983. Water quality standards, which define the quality of water needed to support these activities, have been established and are revised as more is learned about the effects of pollutants on human health and aquatic organisms. The reductions in waste loads needed to achieve these levels of water quality are developed as part of state and local planning processes, and usually include a source-by-source allocation of the required reduction, developed using mathematical water quality models[1].

Monitoring activities, including the collection, evaluation and presentation of water quality data, are a fundamental part of a water pollution control program. Monitoring is not an end in itself; rather the data collected are used to support the other activities of the control agency. Because of the considerable expense of supporting a monitoring program, it is essential to carefully analyze the data requirements and allocate monitoring resources as efficiently as possible.

The basic types of monitoring, ambient, compliance, and biological, used by the control agency in carrying out its program, are designed to detect pollutants in the transport medium, at the source, and at the receptors, respectively.

Ambient monitoring includes fixed station networks, intensive surveys and groundwater monitoring.

The fixed-station network provides data to define histories of the physical and chemical condition of surface water and sediment; trends observed here would typically have to be explained using data from other monitoring activities. Both "representative" sites, to define the general condition of a body of water, and "critical" sites, located at points where standards violations are likely, are included in the fixed-station network.

*Mr. Lewis received his Bachelor degrees in engineering and mathematics from the University of Michigan, and a Masters degree in engineering from MIT. He is an Environmental Systems Analyst on the staff of EPA's Corvallis Environmental Research Laboratory, Corvallis, Oregon. He has published a number of papers on the applications of operations research techniques to environmental problems.

Compliance monitoring consists of permit application, plant process, and compliance schedule verification, checking on the observance of permit pollutant load conditions, and identification of nonfilers.

Checking on the observance of permit conditions includes processing of self-monitoring reports, effluent monitoring to verify the self-monitoring reports and case preparation monitoring.

Biological monitoring is undertaken to determine the suitability of aquatic environments for supporting aquatic organisms, to detect the presence and buildup of toxic substances in aquatic biota, and to provide a basis for updating the eutrophic condition classification of freshwater lakes.

WATER QUALITY STANDARDS

The river basin in which it is desired to control water quality will typically consist of many different types of receiving waters, including streams, lakes, reservoirs and estuaries. Water quality standards have been established to support various uses (e.g. public water supply, recreation, industrial) of each of these categories of water in the basin. Streams are usually further divided into reaches, with the designated use varying as a function of natural quality, and the size of the population and the level of industrial activity adjacent to the stream. Receiving water can be used to assimilate waste as long as these standards are not violated.

There are two basic ways in which wastes can enter water, through direct water use (e.g. withdrawal for industrial use and discharge), or through a combination of normal rainfall runoff and land-use practices adjacent to the receiving waters.

Water quality standards must be met at any time a sample is taken, i.e. the instantaneous or grab sample value of the concentration of a given constituent (usually measured in milligrams per liter or mg/l) must meet the standard for that constituent. The concentration of nitrates, for example, may not exceed $3.0\,mg\,l^{-1}$ in many states. There are other forms of water quality standards, including, in addition to the simple upper limit or not-to-exceed (NTE) example for nitrates, lower limits, or not-less-than thresholds (NLT) (example; dissolved oxygen not-less-than $5\,mg\,l^{-1}$), and those referred to ambient conditions (example: temperature no more than 2° above that due to natural causes).

Ambient water quality (i.e. the concentrations of various constituents in the stream), is stochastic in nature because of variations in the quantity of stream flow and in the waste inputs to the stream (concentration is waste load rate divided by flow). In the absence of other knowledge it is convenient to assume that the instantaneous water quality is normally distributed. Justification for this assumption is usually made on the basis of mathematical tractability and use of the central limit theorem. Under this assumption, the probability of violation, Pv, of a simple-lower-limit (NL) standard is equal to the area under the distribution function to the left of the threshold level.

This distribution does have certain properties which aren't compatible with the physical properties of water quality parameters. Since most of these parameters are concentrations of constituents, there is an absolute, fixed lower limit, usually zero, which means that the negative tail of the normal distribution introduces an error. However, the error is negligible as long as the mean value is large. Reference [3] presents on analysis quantifying the difference between the Gaussian and the Rayleigh distributions; the latter more accurately represents some of the physical properties of water quality parameters.

Note that water quality, $C(t)$ has been assumed to be a random process which can be described by a normal (or some other)) probability distribution at time t. It may be more appropriate to consider water quality, $C(t)$, to consist of both deterministic components (definite amplitudes, phases and frequencies) and random components which are distributed about the mean values of the deterministic variable. This is considered briefly in [3], where it is shown that the normal distribution remains a sufficient representation of $C(t)$ for purposes of developing monitoring network priority procedures.

While the water quality standards discussed above could be applied to effluent discharge (plant outfalls) as well as receiving water, they would be much too strict on the one hand (the substantial dilution achieved in mixing with receiving water would be ignored), and easily circumvented on the other (by adding more dilution water in the plant). Since the real concern is the total load of pollutants the stream system must assimilate, effluent standards are usually

written in terms of allowable mass loading per unit time, i.e., instantaneous concentration limits are replaced by a pollutant load which can be discharged within a given time period. Note that these load limitations will always result in a not-to-be-exceeded (NTE) standard. In writing permits for point sources, the usual practice is to specify the maximum allowable load in a 24-hr period, and the average load for any consecutive 30-day period. Compliance monitoring activity is focused on the 24-hr limit, because the sampling period required to establish compliance or non-compliance is much shorter than that for the 30-day average limitation.

As with water quality concentration variations, there has been only limited progress in statistically characterizing effluent loadings. From what work has been done, the normal and lognormal distributions have been found to adequately model the statistics of daily average effluent loadings. To determine which distribution fits the data, a simple plotting technique (using normal probability paper) is proposed in [4]. This approach gives a simple visual test of the goodness of fit, expecially for large values of the constituent where a violation is likely to occur.

3. MONITORING RESOURCE ALLOCATION

The measure of effectiveness of the monitoring system, i.e. the objective function, can be defined in a variety of ways. One might look at the environmental damage being caused by all of the pollutant discharges and/or water quality violations in the basin, and seek a monitoring system which will minimize that damage. If the damage from all pollutants is expressed in dollars, this approach would automatically account for differences in damage due to, say, a major discharge of a toxic substance (e.g. mercury) and a minor violation of a dissolved solids standard. However, the problem of quantifying damages and expressing them in economic terms is a considerable one, and only limited progress has been made.

An alternative to using explicit damage functions is to provide a weighting scheme which would allow the user to specify the relative importance of violation of standards for various constituents. If all weights are set to unity, all pollutants are treated equally and the minimization of environmental damage reduces to the minimization over the basin of the probability of undetected violations, or, equivalently, the maximization of the probability of detected violations. In this paper, two monitoring resource allocation techniques are discussed in terms of this more restricted definition for the objective function: a procedure for establishing an ambient network (based on maximization of the probability of detected violations) and an effluent monitoring procedure (minimization of probability of non-detected violations).

In both cases the objective function is written as the sum of violation probabilities for individual segments or sources; the summation of probabilities of violation in stream segments is used in the ambient and the summation of probabilities of non-detection of violations at outfalls is used in the effluent procedure. The actual value of this objective function is not usually significant; rather, the concern is that monitoring resources be allocated such that the rate of change (positive or negative, depending upon the objective function definition) is greatest. Also, in both procedures a sequential search of the steepest descent type is used, i.e. the gradient of the objective function is computed and monitoring resources are allocated to move in the direction of maximum rate of change.

The integer nature of the resources to be allocated (plant visits or water quality samplers) leads to a discrete form of the steepest descent technique known as maximum marginal return. Also, the objective function turns out to be a separable function,

$$f\{x_1, x_2, x_3, \ldots, x_n\} = f_1\{x_1\} + f_2\{x_2\} + \cdots + f_n\{x_n\}$$

making calculation of the gradient (or, in this discrete case, the marginal return as the resource expenditure is incremented by one) somewhat easier. In fact, for the effluent monitoring procedure, a complete table of marginal returns for all reasonable values for the number of plant visits to each source can be computed at the outset. As will be seen below, the calculation of this table of values and arrangement in the order of descending magnitude is the core of the effluent monitoring priority procedure. The objective function for the ambient procedure, on the other hand, can only be written in separable form by a transformation of variables. Since this must be done after each new sampling location is selected, the procedure is truly sequential. One further

contrast between the two problem statements: in the ambient problem the summation of probabilities of detected violations over N river segments is maximized by placing samplers to add the next largest term to the summation, while in the effluent procedure the number of terms remains constant (equal to the number of outfalls), and plant visits are allocated to reduce the individual probabilities of non-detection of a violation, such that the maximum reduction in the summation is achieved.

Constraints may exist in the form of regulations requiring, e.g. that every principal outfall be sampled once per year or that the fixed ambient network must have certain representative sites which are placed without regard to the probability of a water quality standard violation. These constraints are assumed to be handled in the procedures discussed here by dividing the budget into fixed and discretionary fractions, with the fixed budget covering required activities and the discretionary budget allocated using these priority procedures. Note that the fixed total budget constraint is easily handled with this incremental approach.

All violation probabilities are computed assuming no evasive action on the part of the polluters, i.e. that they have no advance knowledge of when samples will be taken. Also, both of these procedures are preliminary in that equipment accuracies, reliability, accessibility, etc., are not included in the analysis. This means, among other things, that there is no consideration given to the occurence of Type I and Type II errors, i.e. if there is a violation the equipment is assumed to detect it with no error.

Finally, both procedures have been designed with the realities of data availability in mind.

4. AMBIENT NETWORK OPTIMIZATION

The objective of the ambient monitoring network design procedure developed by Beckers *et al.* [3] is the maximization of the total probability of detection of violations over the entire river basin. The interface between the system being monitored, i.e. the river basin, and the basic characteristics of the monitoring network (specified in terms of approximate sampling location and frequency) is the concern of this design stage. In terms of the earlier discussion of monitoring functions, this procedure optimizes the fixed-station portion of the ambient water quality monitoring category. Further, representative sites, i.e. those necessary to establish water quality background levels, are assumed to have been established, so that the problem becomes one of getting the greatest return from the discretionary portion of the budget available to establish the critical stations.

The basic approach is to look at the river basin from a macroscopic point of view, and replace its three-dimensional physical structure by a network of straight lines representing rivers and streams, and point loads representing clustered outfalls in urban areas. This viewpoint is necessary because of the scope of the network being designed; it must be able to determine the overall status of the entire river basin and major stretches of river within the basin. The temporal scale also reflects the basin-wide viewpoint, with time-average inputs used throughout.

Dividing this macroscopic representation of the river basin into segments is an important step in the procedure, since decisions as to which of these segments to monitor provide the basis for maximizing the objective function, i.e. the probability of detecting violators. Segmentation requires the identification of physical characteristics in or along the stream which cause changes in stream conditions, i.e. either water quality or flow regime. Items which call for establishing a new segment (typically, every 1–10 miles) include clustered outfalls, confluence with a tributary or major changes in the stream bed. In the demonstration of the design procedure on the Wabash River discussed below, there are 33 segments in 482 miles of river.

After the segments have been established, priority ranking is determined for each one using two separate calculations. The first requires the computation of the probability of a violation within the segment, while the second involves making adjustments to this initial segment priority on the basis of knowledge of violations in adjacent segments (the adjusted value is called the ultimate segment priority).

The probability of violation, Pv, of a simple lower bound on concentration (a not-less-than, or NLT standard), Ct, of a given constituent,

$$Pv = P\{C < Ct\} = \int_{-\infty}^{Ct} P\{c\}\,dc = \frac{1}{\sqrt{(2\pi)}\sigma_c} \int_{-\infty}^{Ct} \exp\left[-\frac{\{C-\bar{C}\}^2}{2\sigma_c^2}\right] dc,$$

where it is assumed that the variation in water quality (normal with S.D., σ_c, mean $\bar{C}$) is caused by hydrology, with the waste load described in terms of the aggregate discharged to this stream segment and assumed to be constant over the system duration (several months or more). Since only a rank ordering is desired, the absolute value of the probability of violation isn't needed. It can be shown[3] that the difference begween the expected worst case (for this example, the minimum) mean water quality in the segment, $\bar{C}m$, and the standard, Ct, measured in terms of the S.D. of water quality, σ_c, is sufficient, i.e.

$$P_i = \frac{Ct - \bar{C}m}{\sigma_c}$$

can be used in place of Pv and the evaluation of the integral can be avoided (NLT standard assumed).

Information on the quality of segments up and downstream from the segment under consideration may change the initial priority significantly. If it is known through monitoring that there is a violation of, say, a standard for a conservative parameter (i.e. one which doesn't decay with time) in the segment immediately upstream (segment $N-1$) and no significant new flows are added to dilute the waste load, it is already known that there is likely to be a violation in segment N, and the additional information which could be obtained by monitoring this segment, as reflected in its priority level, should be reduced. It is easily seen that because of the number of segments and the fact that any given segment might be affected by all other segments calculating this reduction would be a formidable job (theoretically there will be $(n/2)$ cross correlations to calculate, each one requiring data which is difficult to get).

The technique used to reduce the computational load to manageable levels is to assume superposition of source impacts on water quality, i.e. that the downstream effect of any single source can be isolated from that of others influencing the concentration at that point. An indicator of the cross-correlation between adjacent segments i and j ($i = j-1$), where the jth segment is under consideration for monitoring coverage, can be computed

$$R_{ij} = \frac{\bar{C}m'}{\bar{C}m},$$

where $\bar{C}m'$ is that portion due to the upstream segment.

Further,

$$R_{ik} = R_{ij}R_{jk} \quad i = j-1 = k-2.$$

This relationship holds exactly for conservative water quality parameters, and is sufficiently accurate for other parameters for use in these approximate priority calculations. It is now possible to compute a maximum expected mean concentration with upstream influences eliminated,

$$\bar{C}m^* = \bar{C}m\{1 - R_{iq}\}$$

and the ultimate segment priority becomes

$$P\uparrow = \frac{Ct - \bar{C}m^*}{\sigma_c}.$$

These equations state that information from segment q "flows" to the other segments to the extent that these segments are correlated with q. In implementing a network of samplers on a river basin with N segments, the first station is placed on the basis of initial segment priorities (since no stations are yet in place, inter-segment correlations have no impact). For the next step in implementation of the network, it is necessary to correct the priority rankings for the remaining $N-1$ segments, based on the extent to which segment q, where the first sampler was placed, gives information on these segments. Thus, the following $N-1$ correlations are

computed,

$$R_{iq} = \begin{cases} R_{i,i+1} \cdot R_{i+1,i+2} \cdot \ldots \cdot R_{q-1,q} & i < q \\ R_{q-1,q} \cdot R_{q-2,q-1} \cdot \ldots \cdot R_{i-1,i-2} & i > q \end{cases}$$

and $R_{iq} = 1$ if $i = q$. Using the equations for $P\dagger$, the $N-1$ ultimate segment priorities may now be recomputed and the next station placed. It is shown in [3] that location of the sampler within the segment is easily determined given the nature of the non-conservative parameter (sampling for conservative parameters can be done anywhere within the segment).

The optimization approach used is a discrete form of the steepest descent technique called maximum marginal return. For the ambient network, the location of more-or-less permanently installed samplers are being determined (as opposed to plant visit scheduling under the effluent procedure), and a zero-one type decision is involved—either place the sampler or not. The marginal return at any stage of implementation is then simply the ultimate segment priority, and selection of the segment with the largest ultimate priority gives the greatest increase in the objective function (i.e. maximizes the marginal return).

The procedure for locating samplers which has been described gives a preferred location at each stage of implementation of the network. Differences in the probability of detection, given that a violation exists, have not yet been considered. The problem of selecting the sampling frequency, which brings the probability of detection into the problem, has been approached in [3] by defining a priority measure which describes the increasing relative effectiveness of sampling as the frequency is increased. A preferred frequency is not indicated, but when the priority measure is combined with a cost function describing costs of sampling (which increase with frequency), a cost-effectiveness analysis can be conducted. The priority measure is defined

$$M = \frac{\text{Expected Number of Violations Detected}}{\text{Expected Number of Violations}}.$$

An expression for M in terms of the sampling internal, Δ (the reciprocal of the sampling frequency), the expected average duration of a violation, T_1, and the expected average interval between violations, T_0 is developed in [3] as

$$M\{\Delta, T_0, T_1\} = \frac{T_0}{\Delta}\left[C_0 \exp\left\{\frac{-\Delta}{T_0}\right\} + C_1 \exp\left\{\frac{-\Delta}{T_1}\right\}\right],$$

where C_0 and C_1 are functions of T_0 and T_1 (not Δ), and $\Delta < T_0$ has been assumed. This relationship was derived based on the probability of a violation event, assuming the events to be Poisson processes. For $\Delta < T_0 + T_1$, the case of most interest (those in which a high percentage of violations are to be detected, i.e. $M > 50\%$) this expression becomes

$$M\{\Delta, T_0, T_1\} \cong M^*\{\Delta, T_0, T_1\}$$

$$\cong \frac{T_0}{\Delta}\left[\frac{\exp\left\{\frac{-\Delta}{T_0}\right\} - \exp\left\{\frac{-\Delta}{T_1}\right\}}{\left\{\frac{T_0}{T_1}\right\} - 1}\right].$$

This expression is plotted in Fig. 1 as a function of the relative sampling interval, Δ/T_1 for several values of T_0/T_1. As Δ/T_1 becomes small, and the probability of at least one sample being taken during each violation is close to one, M^* approaches unity as expected. Also, as T_0/T_1 increases, corresponding to a small probability of violation, M^* becomes a function only of Δ/T_1 (i.e. independent of T_0).

The major features of the quantitative preliminary design procedure are demonstrated in [3] on the Wabash River in Indiana and Illinois. A brief description of a portion of this demonstration is included here as an illustration of the application of the procedure.

The first task in preliminary design of water quality surveillance systems is primarily

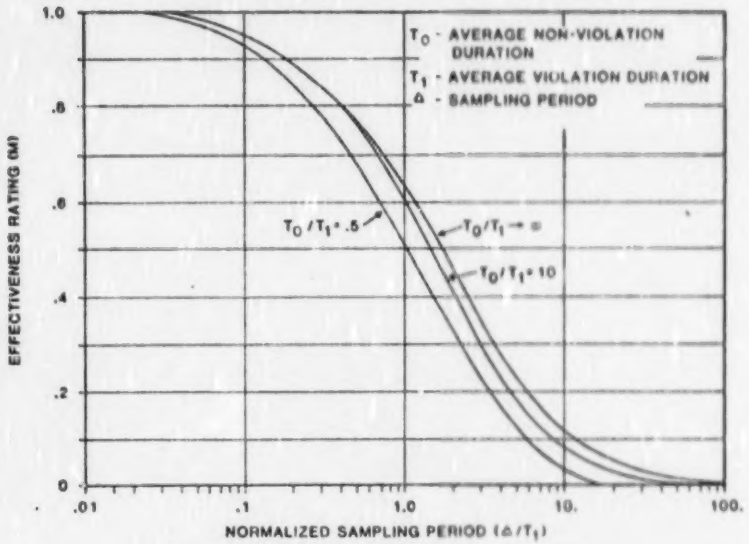

Fig. 1. Temporal effectiveness rating as a function of T_1, T_0 and Δ [3].

organizational. Included in the task are map gathering, standards identification, data organization, and selection of constraints that will guide subsequent activity. Data on the Wabash River used in the demonstration included:

(1) Ambient water quality data from the EPA water quality data files for a number of stations in the Wabash River Basin.

(2) Inventories of municipal and industrial wastewater discharges.

(3) Daily stream flows and stream flow cross-sections from various publications of the U.S. Geological Survey (USGS).

(4) Charts and maps, including USGS maps to provide an overview of the basin, and U.S. Army Corps of Engineers navigation charts showing river mileage, bottom contours and river grade.

The so-called "basin subset," those streams in the river basin to be included in the monitoring network, is specified to be the Wabash mainstem (Fig. 2). A system duration of two months, covering the August/September low flow period of the year, was selected. The data on the Wabash indicate few problems on the river during the remainder of the year. The water quality parameters selected for system design consideration were biochemical oxygen demand and dissolved oxygen (BOD–DO), which are representative of nonconservative, coupled parameters, and are parameters of great interest in water quality evaluation.

The results of source grouping and stream segmentation is presented in Table 1 and diagrammed in Fig. 2. The tolerance lengths, on which source grouping and segmentation is based, vary from fractions of a mile to over 5 miles. A number of single and grouped discharges into the Wabash were ignored during the preliminary design process.

The standard deviations of water quality, σ_c, were estimated from historical water quality records. There were insufficient data on the variability of effluent sources to permit use of mathematical models in characterizing parameter variability. The standard threshold (CI) used for BOD–DO (5 mg l^{-1}) is the value set by Indiana for use of the waters to support aquatic life.

In the first case, it is assumed of interest to establish a surveillance system to detect BOD–DO violations on the Wabash. Considering only priority as a criterion for station implementation, the first station would be in segment 25, below Terre Haute, where the initial segment priority is +0.178 (column 14 in Table 2). Computation of the ultimate segment priority, based on a station in segment 25, gives the values in the column labeled P_i^*. It can be seen that the priorities in segments 24 and 26 are radically lowered, while those in segments 15 through 23 and 27–33 are

Fig. 2. Segmentation of wabash river [3].

somewhat lower. Upstream of segment 15 there is no change, due primarily to the low correlation between segment 14 and 15 (the column labeled R gives the correlation with the upstream segment, Rij for $i = j - 1$.) Following computation of ultimate segment priority, segment 2 has the highest ultimate priority of those remaining (+0.122). This may be compared with the ultimate segment priority of segment 24 (−2.30). The initial segment priority in segment 24 was 0.167, nearly equal to that of segment 25. Thus, consideration of intersegment correlation is an important feature of the system design process.

The procedure for determining time-sampling ratings for BOD was exercised for two cases of interest on the Wabash River. The cases are the two segments with highest ultimate segment priority, segment 25 and segment 2. Four candidate sampling frequencies are considered: daily, weekly, monthly, and quarterly. The base data are for the period of system duration (August and September) over the years 1961–1965. Table 3 is the computation table for the two demonstration cases. It can be seen that, for some purposes, a weekly sampling rate might be appropriate in segment 25, below Terre Haute, since M is 0.7 under that sampling scheme.

In contrast, a daily sampling rate would be required to obtain a comparable M in the headwaters of the Wabash. A weekly sampling rate at the headwaters yields a very poor $M = 0.17$. The result is intuitively satisfactory; the volume of flow in segment 2 is 0.2% of the flow in segment 25 and subject to greater variability over short periods of time. A higher frequency of observation should be anticipated under those circumstances. The quantitative procedure provides a numerical measure of precisely how much higher the frequency must be to achieve the same results.

This discussion of the ambient monitoring design procedure has focused on the single parameter situation. In any actual network design problem, monitoring of a number of water quality parameters will be desired, and there may be as many preferred station locations and priorities as there are parameters. In a preliminary design sense, it is sufficient to compute an aggregate priority for each segment using the average of these individual segment priorities for

Table 1. Segmentation computation table for Wabash River[3]

Seg. No. (N)	Segment position R.M.I.* Upstream end	Downstream end	x_{ave} (Miles)	Stream depth (ft)	Flow velocity (ft s^{-1})	Stream flow (l s^{-1})	Water temp. (°C)	DO sat. conc.	k_a (l day^{-1})	Discharge characterization Industry—S.I.C. code Municipal—treatment tributary—tributary name	Discharge flow rates (l s^{-1})	Discharge concentrations (mg l^{-1}) BOD	DO Deficit	k_a (l day^{-1})	Seg. No. (N)
1	482	469	0.49	0.5	0.4	85	23	8.4	18	(1) Industrial—2023	$Q_1 = 0.878$	$L_1 = 360$	$D_1 = 0$		
										(2) Industrial—2015	$Q_2 = 8.8$	$L_2 = 72$	$D_2 = 0$		
										(3) Municipal—Secondary	$Q_3 = 9.35$	$L_3 = 20$	$D_3 = 6.4$		
										Grouped Totals	$Q_a = 19.0$	$L_a = 59.6$	$D_a = 3.15$	1.4	1
										Tributary—Beaver Creek	$Q_a = 108$	$L_a = 3.14$	$D_a = 4.6$	1.4	2
2	469	456	1.1	0.5	0.9	193	23	8.4	33	(1) Industrial—2033	$Q_1 = 8.78$	$L_1 = 120$	$D_1 = 0.031$		
3	456	452	0.69	0.5	0.6	195	23	8.4	24	(2) Municipal—Raw	$Q_2 = 7.37$	$L_2 = 320$	$D_2 = 8.4$		
										Grouped Totals	$Q_a = 16.15$	$L_a = 211$	$D_a = 3.84$	1.5	3
⋮										(3) Industrial—3400	$Q_3 = 2.6$	$L_3 = 36$	$D_3 = 1.0$		
										(4) Municipal—Secondary	$Q_4 = 4.4$	$L_4 = 73.2$	$D_4 = 6.1$		
										Grouped Totals	$Q_a = 98.6$	$L_a = 65.6$	$D_a = 3.33$	0.45	27
										(1) Municipal—Primary	$Q_1 = 94$	$L_1 = 106$	$D_1 = 6.0$		
										(2) Industrial—2611	$Q_2 = 119$	$L_2 = 80$	$D_2 = 0.229$		
78	128	122	2.99	3.5	0.80	41,600	26	8.0	1.2	Grouped Totals	$Q_a = 213$	$L_a = 91.3$	$D_a = 2.78$	0.46	28
										Tributary—Embarras	$Q_a = 1700$	$L_a = 3.4$	$D_a = 1.6$	0.41	29
										River					
29	122	115	3.15	6.0	0.75	43,583	25	8.1	0.61	(1) Municipal—Secondary	$Q_1 = 5.8$	$L_1 = 87.2$	$D_1 = 6.1$		
										(2) Municipal—Primary	$Q_2 = 5.3$	$L_2 = 149$	$D_2 = 6.1$		
30	115	95	4.9	7.2	0.60	44,500	25	8.1	0.47	Grouped Totals	$Q_a = 11.1$	$L_a = 117$	$D_a = 6.1$	0.31	30
										(1) Tributary—White River	$Q_1 = 42,500$	$L_1 = 2.44$	$D_1 = 1.43$		
										(2) Industrial—2661	$Q_2 = 19$	$L_2 = 3181$	$D_2 = 0.573$		
										(3) Municipal—Primary	$Q_3 = 53$	$L_3 = 102$	$D_3 = 6.1$		
31	95	89	3.25	8.2	0.70	98,705	25	8.1	0.40	(4) Tributary—Patoka River	$Q_4 = 1133$	$L_4 = 4.4$	$D_4 = 2.7$		
										Grouped Totals	$Q_a = 43,705$	$L_a = 3.99$	$D_a = 1.47$	0.37	31
										Municipal—Secondary	$Q_a = 14$	$L_a = 40$	$D_a = 6.25$	0.21	32
										(1) Tributary—Bonpass Creek	$Q_1 = 57$	$L_1 = 3.14$	$D_1 = 0.423$		
32	89	65	5.7	8.7	0.70	98,749	24	8.25	0.40	(2) Municipal—Primary	$Q_2 = 13$	$L_2 = 147$	$D_2 = 6.25$		
33	65	0	3.22	8.2	0.70	98,819	24	8.25	0.40	Grouped Totals	$Q_a = 70$	$L_a = 30.2$	$D_a = 1.51$	0.37	33

*R.M.I.—River Mile Index, miles from the mouth of the Wabash River.

Table 2. Priority computation table for Wabash River [3]

Seg. No. (N)	Segment position R.M.I. Upstream end	Downstream end	x_E (mile)	L_o mg l^{-1}	D_o mg l^{-1}	x_m mile	D_m mg l^{-1}	$\bar{C}_m$ mg l^{-1}	D_E mg l^{-1}	L_E mg l^{-1}	σ mg l^{-1}	C_s mg l^{-1}	P Priority rating with no impl	R Correlations	D_t°	C_t°	P_t° (impl on segment 25)
1	482	469	13	10.9	0.704	0.434	0.8	6.54	0.057	0.68	0.9	5	−1.71	0	0.8	6.54	−1.71
2	469	456	13	2.43	2.51	0	2.51	4.89	0.031	0.71	0.9	5	0.122	0.02	2.59	4.89	0.122
3	456	452	4	18.2	0.345	1.17	0.965	6.44	0.659	9.9	0.9	5	−1.6	0.04	0.965	6.44	−1.6
4	452	434	18	10.1	0.659	0.292	0.778	6.62	0.064	0.576	0.9	5	−1.8	0.98	0.778	6.62	−1.8
5	434	422	12	7.77	1.34	0	1.34	6.06	0.119	2.11	0.9	5	−1.18	0.05	1.34	6.06	−1.18
6	422	406	16	6.14	0.224	1.34	0.310	6.79	0.092	1.53	0.9	5	−1.99	0.37	0.310	6.79	−1.99
7	406	394	12	2.48	0.527	0	0.527	6.57	0.036	0.73	0.9	5	−1.74	0.08	0.527	6.57	−1.74
8	394	387	7	9.51	0.726	0	0.726	6.52	0.413	6.2	0.9	5	−1.69	0.01	0.726	6.52	−1.69
9	387	375	12	9.4	0.805	0	0.805	6.45	0.322	4.33	0.9	5	−1.61	0.50	0.805	6.45	−1.61
10	375	372	3	6.49	0.760	0	0.760	6.64	0.370	5.40	0.9	5	−1.82	0.27	0.760	6.64	−1.82
11	372	363	9	5.62	0.487	0	0.487	6.91	0.264	3.51	0.9	5	−2.12	0.73	0.487	6.91	−2.12
12	363	354	9	3.68	0.390	0.182	0.405	6.99	0.382	2.05	0.9	5	−2.21	0.61	0.405	6.99	−2.21
13	354	329	25	2.40	1.40	0	1.40	8.20	0.097	0.732	0.9	5	−1.33	0.12	1.40	6.20	−1.33
14	329	322	7	0.878	0.198	0	0.198	7.20	0.096	0.740	0.9	5	−2.44	0.47	0.198	7.20	−2.44
15	322	317	5	1.29	0.861	0	0.861	6.54	0.594	0.748	0.9	5	−1.71	0.07	0.852	6.55	−1.72
16	317	312	5	0.80	0.55	0	0.55	6.85	0.385	0.695	0.9	5	−2.06	0.92	0.51	6.89	−2.10
17	312	287	25	7.16	0.50	14.4	1.19	6.06	1.06	3.66	0.9	5	−1.18	0.31	1.08	6.17	−1.3
18	287	271	16	3.86	1.07	0	1.07	6.18	0.705	2.81	0.9	5	−1.31	0.99	0.770	6.48	−1.64
19	271	257	14	3.17	0.71	11.5	0.904	6.35	0.898	2.26	0.9	5	−1.50	0.92	0.842	6.61	−1.79
20	257	245	12	2.42	1.05	1.28	1.06	6.19	0.964	1.62	0.9	5	−1.32	0.81	0.731	6.52	−1.69
21	245	239	6	1.7	0.95	6.0	1.05	6.20	1.05	1.36	0.9	5	−1.33	0.88	0.651	6.60	−1.78
22	239	230	9	1.45	0.99	6.49	1.03	6.22	1.03	0.982	0.9	5	−1.36	0.98	0.577	6.67	−1.86
23	230	214	16	1.65	1.03	7.63	1.08	6.17	1.05	1.45	0.9	5	−1.30	0.94	0.594	6.66	−1.84
24	214	196	18	5.12	1.12	18	2.25	4.85	2.25	2.80	0.9	5	0.167	0.45	0.023	7.07	−2.30
25	196	172	24	2.84	2.26	0	2.26	4.84	1.95	1.39	0.9	5	0.178	0.99	Impl		
26	172	158	14	1.41	1.95	0	1.95	5.15	0.926	0.973	0.9	5	−0.17	0.99	0.020	7.08	−2.31
27	158	128	30	1.15	0.932	0	0.932	6.17	0.221	0.322	0.9	5	−1.30	0.99	0.019	7.08	−2.31
28	128	122	6	1.43	0.243	6	0.405	6.60	0.405	1.16	0.9	5	−1.78	0.52	0.198	6.80	−2.00
29	122	115	7	1.25	0.605	7	0.649	6.45	0.649	0.989	0.9	5	−1.61	0.72	0.401	6.70	−1.89
30	115	95	20	1.02	0.65	1.04	0.651	6.45	0.573	0.504	0.9	5	−1.61	0.99	0.410	6.69	−1.88
31	95	89	6	2.00	0.912	6	1.06	6.04	1.06	1.65	0.9	5	−1.16	0.49	0.869	6.23	−1.37
32	89	65	24	1.67	1.07	0	1.07	6.03	0.848	1.08	0.9	5	−1.17	0.95	0.888	6.36	−1.51
33	65	0	65	1.10	0.849	4.96	0.851	6.40	0.345	0.134	0.9	5	−1.56	0.98	0.706	6.54	−1.71

Table 3. Computation table for temporal priority rating for Wabash Mainstem[3]

N'	Implementation sequence No.	N	Segment No.	Q(n−1)	Stream flow (l s⁻¹)	Qe	Effluent group flow (l s⁻¹)	Cse	Saturation DO conc. (mg l⁻¹)	Cm	Maximum DO cond. (mg l⁻¹)	Cr	Threshold DO conc. (mg l⁻¹)	Dm	Background DO deficit (mg l⁻¹)	Qr	Threshold stream flow (CFS)		Stream flow station	ΔQ	Stream flow adjustment (CFS)	To	Nonviolation duration (days)	Tv	Violation duration (days)	Δ	Sampling period (days)	M	Priority rating	Δ	Sampling period (days)	M	Priority rating	Δ	Sampling period (days)	M	Priority rating
1	1	25	25	36307	23.7	8.1	4.84	5.0	1.0	1380	3.3		3415	0	0	46.8	13	1	0.96	7	0.70	30	0.33	90	0.10
2	2		2	85	108	8.4	4.89	5.0	1.0				3225	0	0	25.1	1.2	1	0.66	7	0.17	30	0.03	90	<0.01
3																																			
4																																			
5																																			
6																																			
7																																			
8																																			
9																																			

each parameter in the segment. Since correlations between segments are computed on a parameter-by-parameter basis, it is necessary to compute new correlations and a new average each time a new station is implemented.

5. EFFLUENT NETWORK OPTIMIZATION

The effluent monitoring priority procedure, developed by Cohen *et al.*[4] is a preliminary design tool in the same sense as the ambient procedure described in Section 4.0 (i.e. no detailed consideration is given to instrumentation accuracy, reliability, etc.). The system to be monitored is already discrete, since there is a given number of point sources in the basin, each with a permit containing a specification of the allowable daily maximum and monthly average load for several constituents which can be discharged to surface waters.

Again the resources to be allocated are discrete additional visits to sample an outfall and, although the initial specification of the objective function is somewhat more elaborate, the solution technique will be seen to be equivalent to the discrete steepest descent technique (maximum marginal return) used in the ambient system design procedure.

One important difference between the ambient and effluent design problem is the time scale over which the system parameters are assumed to remain constant (the "system duration" used in [3]) and the way in which the two procedures handle changes in system parameters.

The effluent system is concerned with day-to-day fluctuations in individual outfalls. These fluctuations are described by statistical distribution functions which themselves have short-term variations due to changes in production levels, production process modifications, etc. Because of the need to follow these variations in individual source statistics, the characteristic time scale associated with the effluent system is much shorter (say, one month) than that for the ambient systems (several months or more). To update the statistical parameters, a Bayesian approach is used, with new data obtained from self-monitoring reports and prior compliance monitoring data. This Bayesian update calculation is an integral part of the effluent monitoring design procedure, and is appropriate to an allocation system which can respond quickly to changes in priority, since it is scheduling effluent sampling visits to individual plants. The ambient design procedure, which allocates more-or-less permanently installed fixed sampling stations rather than plant visits, doesn't include any update capability *per se*; if there are changes in the system parameters the procedure must be applied again from the beginning. One important feature of both the initial statistical characterization of effluent levels and the subsequent update calculations is the use of only that data which is immediately and routinely available to the user.

The objective function developed in [4] is based on the notion that effluent monitoring resources should be allocated in such a way that the environmental damage (or cost) to receiving waters caused by pollutant discharge is minimized. Further, the assumption is made that the environmental cost of a detected violation is zero (i.e. detection leads to some corrective action), leading to the statement of the problem as the minimization of the environmental cost of undetected permit condition violations.

For a single outfall, the total environmental cost incurred when the outfall is sampled S times (one day for each sample, since compliance with the daily maximum is tested) during a period consisting of N intervals (typically 30 intervals of one day each) can be written using the total probability law. That is, the cost is the expected damage given that a violation was not detected times the probability that the violation was not observed plus the expected damage given that a violation was detected times the probability that a violation was detected.

Mathematically,

$$C_i\{S_i\} = \frac{1}{N} E\left\{\sum_{k=1}^{N} d_i\{k\} \middle| \bar{V}_i\right\} P\{\bar{V}_i | S_i\} + \frac{1}{N} E\left\{\sum_{k=1}^{N} d_i\{k\} \middle| V_i\right\} P\{V_i | S_i\},$$

where $C_i(S_i)$ = total cost, ith source when sampled S times
$d_i(k)$ = damage incurred due to the ith source during kth interval
$\bar{V}_i$ = event that no violation is observed when sampling the source.

Because the environmental cost of a detected violation is assumed to be zero, the second term on the right-hand side above is zero. Also, the development of the damage functions $d_i(k)$ is rather involved and wont't be reported here (see [4] for these functions). For the purposes of comparing analytical techniques used in developing the ambient and effluent design procedures, $d_i(k)$ is set to unity for all i and all k. With these changes the total cost becomes

$$\{C_i\{S_i\} = P\{\bar{V}_i|S_i\}.$$

As discussed in Section 2.1, the probability of no violation of a simple threshold effluent standard given a single sample of source i is β_i. For S_i samples,

$$P\{\bar{V}_i|S_i\} = \beta_i^{S_i}$$

assuming independence between concentrations at various sampling times. The total basin cost becomes

$$C = \sum_{i=1}^{N_s} C_i\{S_i\} = \sum_{i=1}^{N_s} \beta_i^{S_i},$$

where N_s = number of effluent sources in the basin.

To minimize this separable function using a steepest descent technique, it is only necessary to compute rates of change of individual terms and then select that term with the largest rate. As with the ambient design procedure, resources are allocated in discrete amounts (sampling visits) and the gradient becomes a series of marginal changes

$$\mu_i = \frac{\beta_i^{S_i-1} - \beta_i^{S_i}}{r_i},$$

where r_i is the cost of collecting the additional sample (r_i's must all be of equal size or each one must be very much smaller than the total budget to yield nearly optimum allocations). In calculating these marginal changes for each term, allocations of more than one visit should be considered, since return visits to sources with low probabilities of no violation will normally be required in the optimal allocation. Arranging the μ_i in decreasing order then gives the prioritized list of sources to be sampled, allowing the control agency to proceed as far down the list as their monitoring budget will allow, with the assurance that the greatest possible return has been achieved. Selecting an outfall for allocation of the next visit doesn't require a complete recalculation of marginal returns, as is required at each stage of the ambient design procedure, because the outfalls are independent of one another.

The Bayesian update calculation has been made an integral part of the effluent design procedure because of the need for the procedure to respond to changes in source statistics. At the end of each monitoring period (e.g. every month) the user recalculates the mean and variance for each source using self-monitoring data which has been submitted by each source, as well as the monitoring data collected during the prior monitoring period. The user also has the problem of initially characterizing the sources before any data is available. Techniques are developed in [4] which allow the user to estimate these statistical parameters from information and or data which is available to him—daily maximum and monthly mean from the conditions in the permit are used for the initial statistical description, and self-monitoring and compliance data are used for updating the statistical description. The block diagram in Fig. 3 shows how these calculations proceed.

Multiple constituents at each source can be handled by the current design procedure under two situations: independent constituents

$$\beta_i = \prod_j \beta_{ij},$$

or completely correlated

$$\beta_i = \min_j \beta_{ij}.$$

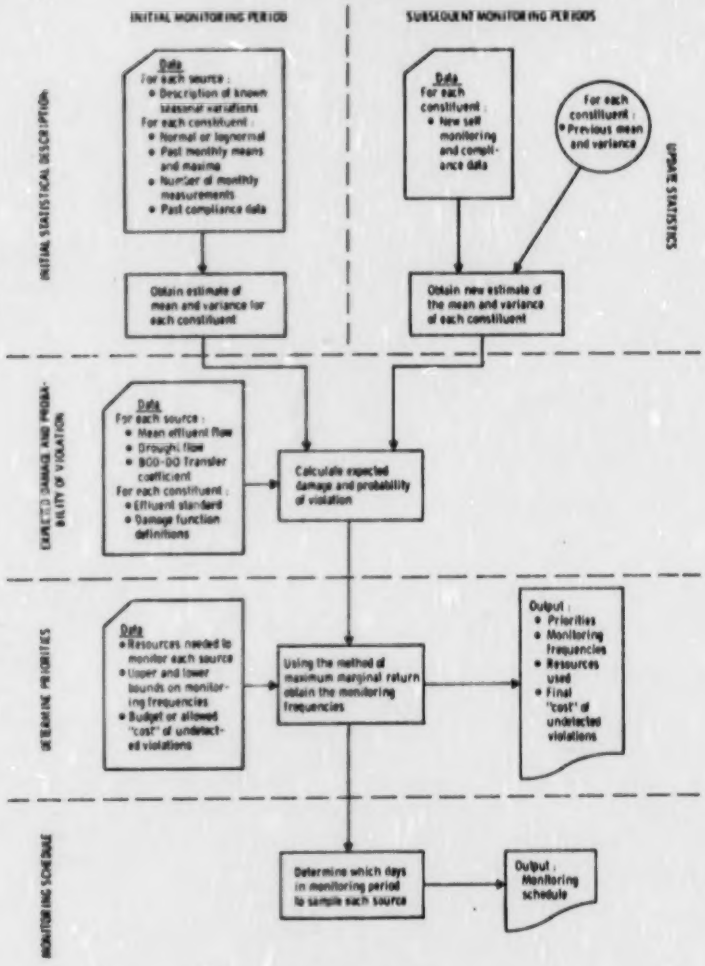
INITIAL MONITORING PERIOD
SUBSEQUENT MONITORING PERIODS
INITIAL STATISTICAL DESCRIPTION
UPDATE STATISTICS
Data
For each source :
• Description of known seasonal variations
For each constituent :
• Normal or lognormal
• Past monthly means and maxima
• Number of monthly measurements
• Past compliance data
Obtain estimate of mean and variance for each constituent
Data
For each constituent :
• New self monitoring and compliance data
For each constituent :
• Previous mean and variance
Obtain new estimate of the mean and variance of each constituent
EXPECTED DAMAGE AND PROBABILITY OF VIOLATION
Data
For each source :
• Mean effluent flow
• Drought flow
• BOD-DO Transfer coefficient
For each constituent :
• Effluent standard
• Damage function definitions
Calculate expected damage and probability of violation
DETERMINE PRIORITIES
Data
• Resources needed to monitor each source
• Upper and lower bounds on monitoring frequencies
• Budget or allowed "cost" of undetected violations
Using the method of maximum marginal return obtain the monitoring frequencies
Output :
• Priorities
• Monitoring frequencies
• Resources used
• Final "cost" of undetected violations
MONITORING SCHEDULE
Determine which days in monitoring period to sample each source
Output :
Monitoring schedule

Fig. 3. Effluent monitoring priority procedure [4].

As an illustration of the use of the priority procedure a simplified example is presented in [4]. Initially, it is assumed that there are four sources to be monitored, each having four months of self-monitoring data available from which to obtain the initial statistics. The initial self monitoring data assumed and the initial statistics obtained from this data are shown in Table 4 for two of the four sources. The estimated mean and estimated standard deviation are the monthly estimates using the techniques developed in [4]. Table 4 also shows how the estimates of the mean and standard deviation are sequentially updated as the monthly estimates are combined to obtain the estimates to be used in the priority procedure. The expected damage (d_i used as developed in [4]), and probability of violation (d_i set to one) obtained from the data are shown in Table 5, along with the estimated source flow and the stream flow. Source 3 is a large sewage treatment plant. From the table the impact of BOD and phosphates is large; however, the standards are also large and therefore the probability of violation for the parameters is small. Source 4 has a relatively small impact on the stream (i.e. small expected damage); however, the standards have been set so that the probability of violation is very large.

The priority listing of the sources is shown in Table 6. The dollar figures are based on rough estimates ([4]) of the cost of sampling each of the constituents. For the purposes of this example, it was assumed that the sources could be sampled between 0 and 10 times. From the table, one sees that Sources 1 and 3 should be sampled with higher priority than sources 2 and 4. This is due to the much larger expected damage from the former sources. Source 4 appears relatively early in the list, but most of the samples have low priority. This is because the

Table 4. Statistical data for the effluent monitoring priority procedure[4]

Source monitoring data

Source 1

Month	Mean source flow (ml day^{-1})	Parameter: pH max eff. standard: 9 distribution: normal Mean	Max.	Sample size	Parameter: pH min eff. standard: 6 distribution: normal Mean	Min.	Sample size	Parameter: lead eff. standard: 2 kg distribution: normal Mean (kg)	Max. (kg)	Sample size
1	0.90	8.5	10.6	20	8.5	6.0	20	0.41	1.0	20
2	1.10	7.6	9.0	22	7.6	5.4	22	1.08	1.7	22
3	1.20	8.3	9.8	22	8.3	6.4	22	1.09	6.3	22
4	0.85	8.1	9.5	20	8.1	6.4	20	0.52	1.8	22

Source 2

Month	Mean source flow (ml day^{-1})	Parameter: chromium eff. standard: 0.48 kg distribution: normal Mean (kg)	Max. (kg)	Sample size	Parameter: copper eff. standard: 1.5 kg distribution: lognormal Mean (kg)	Max. (kg)	Sample size	Parameter: fluoride eff. standard: 30 kg distribution: normal Mean (kg)	Max. (kg)	Sample size
1	0.80	0.216	0.808	18	0.524	1.89	18	24.4	31.4	18
2	0.78	0.313	0.867	19	0.374	1.87	19	25.4	31.9	19
3	0.87	0.214	0.620	21	0.364	1.25	22	24.7	31.0	22
4	0.85	0.132	0.255	14	0.110	0.42	14	14.0	31.0	11

Initial statistics

Source 1

Month	Parameter: pH max. distribution: normal Est. mean	Est. S.D.	Updated mean	Updated S.D.	Parameter: pH min. distribution: normal Est. mean	Est. S.D.	Updated mean	Updated S.D.	Parameter: lead distribution: normal Est. mean (kg)	Est. S.D. (kg)	Updated mean (kg)	Updated S.D. (kg)
1	8.5	1.12	—	—	8.5	1.33	—	—	0.41	0.31	—	—
2	7.6	0.73	8.03	1.06	7.6	1.15	8.03	1.33	1.08	0.32	0.76	0.51
3	8.3	0.78	8.12	0.98	8.3	0.99	8.12	1.22	1.09	2.72	0.87	1.62
4	8.1	0.74	8.12	0.92	8.1	0.90	8.12	1.14	0.515	0.67	0.78	1.45

Source 2

Month	Parameter: chromium distribution: normal Est. mean (kg)	Est. S.D. (kg)	Updated mean (kg)	Updated S.D. (kg)	Parameter: copper distribution: lognormal Est. mean (log kg)	Est. S.D. (log kg)	Updated mean (log kg)	Updated S.D. (log kg)	Parameter: fluoride distribution: normal Est. mean (kg)	Est. S.D. (kg)	Updated mean (kg)	Updated S.D. (kg)
1	0.216	0.321	—	—	−0.437	0.369	—	—	24.4	3.79	—	—
2	0.313	0.297	0.266	0.308	−0.685	0.474	−0.565	0.443	25.4	3.49	24.9	3.62
3	0.214	0.214	0.247	0.277	−0.570	0.337	−0.567	0.403	24.7	3.29	24.8	3.46
4	0.132	0.070	0.218	0.246	−1.146	0.404	−0.771	0.502	24.0	4.17	24.6	3.61

Table 5. Expected damage and probability of violation[4]

Source	Pipe	Est. source flow (ml day^{-1})	Stream flow (ml day^{-1})	Parameter	Expected damage D_0	Probability of no violation, P_0 (%)	Expected damage for source, c_i	Probability of no biolation for source, p_i (%)
1	1	0.961	100	pH	0.29	80.0		
				Lead	1.60	80.0	1.60	64.0
2	1	0.845	320	Chromium	0.08	82.6		
				Copper	0.12	96.1		
				Fluoride	0.00	93.1	0.12	74.0
3	1	108	525	BOD_5	3.23	100.0		
				Phosphate	3.64	97.6		
				Suspended solids	0.37	87.8	3.64	85.6
4	1	0.297	300	Phosphates	—	100.0		
				Suspended solids	—	51.8		
	2	1.016		Phosphates	0.29	54.4		
				Suspended solids	0.03	46.0	0.29	13.0

Table 6. Priority listing[4]

Priority	Source sampled	Marginal return X100	Cost of undetected violations	Resources required
1	1	0.10774492	5.07571	535.50
2	3	0.09326524	4.55342	1095.50
3	3	0.07959130	4.10603	1655.50
4	1	0.06899248	3.73658	2191.00
5	3	0.06843515	3.35334	2751.00
6	3	0.05862177	3.02506	3311.00
7	3	0.05021559	2.74385	3871.00
8	4	0.04526206	2.49264	4426.00
9	1	0.04417806	2.25607	4961.50
10	3	0.04301484	2.01519	5521.50
11	3	0.03684665	1.80885	6081.50
12	3	0.03156296	1.63209	6641.50
13	1	0.02828861	1.48601	7177.00
14	3	0.02703693	1.32920	7737.00
15	3	0.02315992	1.19951	8297.00
16	1	0.01811409	1.10251	9832.50
17	1	0.01159902	1.04039	9368.00
18	1	0.00742722	1.00062	9903.50
19	4	0.00590254	0.96786	10458.50
20	2	0.00556719	0.93735	11006.50
21	1	0.00475588	0.91188	11542.00
22	2	0.00412025	0.88931	12090.00
23	2	0.00304938	0.87260	12638.00
24	1	0.00304534	0.85629	13173.50
25	2	0.00225683	0.84392	13721.50
26	1	0.00195003	0.83348	14257.00
27	2	0.00167027	0.82432	14805.00
28	2	0.00123616	0.81755	15353.00
29	2	0.00091488	0.81254	15901.00
30	4	0.00076974	0.80826	16456.00
31	2	0.00067710	0.80455	17004.00
32	2	0.00050112	0.80181	1755.200
33	2	0.00037087	0.79978	18180.00
34	4	0.00010038	0.79922	18655.00
35	4	0.00001309	0.79915	19210.00
36	4	0.00000171	0.79914	19765.00
37	4	0.00000022	0.79914	20320.00
38	4	0.00000003	0.79914	20875.00
39	4	0.0000000	0.79914	21430.00
40	4	0.00000000	0.79914	21985.00

probability of violation is very large and therefore the chances are that the source will be caught in violation after one or two visits. Further sampling is therefore not necessary. Source 2 has a small expected damage and a fairly large probability of no violation resulting in a low sampling priority. Tabie 6 also gives the marginal return, "cost" of undetected violations and resources used. The marginal returns are decreasing (the list has been ordered in just this manner). The "cost" of undetected violations is decreasing, and the resources required are increasing as more sources are sampled.

If only, say, $10,000 were available for monitoring, then only the sources with priority 1–18 would be monitored. If, on the other hand, a maximum allowed "cost" of undetected violations of, say, 1.0 were specified, then sources with priorities 1 through 19 would be sample. The priority list also shows when the return from monitoring (i.e. the marginal return) starts becoming negligible; the return, in this case, for monitoring more than, say, 25 sources is very small.

6. ADDITIONAL WORK

The procedures which have been described were developed under different Federal laws with fundamentally different approaches-to water pollution control. Also, the effluent monitoring procedure, developed under current legislation, was completed before the detailed monitoring regulations were issued.

Regulations pertaining to monitoring have been now issued[1, 5, 6], and a general description

of state water monitoring programs has been published[2]. The first priority need is to consider the overall monitoring problem, bring in the other monitoring functions (e.g. biological) and integrate the ambient and effluent priority procedure within an overall monitoring systems framework. Outputs from the ambient network, for example, could be used to update the statistics of all sources in the vicinity of a water quality violation. Another useful input to the effluent monitoring priority procedure would be to use the results of intensive surveys, possibly tied to the update block on a priority interrupt basis rather than the routine update interval used in the current procedure.

A subsequent ambient monitoring report by Beckers and Chamberlain[7] goes beyond the preliminary procedures described above and addresses some of the engineering aspects of ambient network design, (e.g. equipment cost, accuracy, reliability, maintainability) and computerizes the design procedure. This level of design needs to be undertaken for the effluent monitoring system, and should include techniques for scheduling monitoring visits taking into account manpower, equipment and geographical factors. Further analysis of the statistical properties of effluent levels, including study of the correlation between constituents, is also needed.

Finally, these techniques should be demonstrated under field conditions. Limited demonstrations to date on the ambient and effluent procedures show that these techniques work and do result in more efficient networks, see [3 and 4]. However, better estimates of the payoff which can be realized using these techniques and the costs of applying them are needed before any conclusions can be drawn regarding the utility of operations research in optimizing monitoring programs, given real-world problems of data availability and user expertise.

REFERENCES

1. Water Quality Management Basin Plans, 40 CFR Part 130, Federal Register, Vol. 39, No. 107 (3 June 1974).
2. U.S. Environmental Protection Agency, Model State Water Monitoring Program, EPA-440/9-74-002 (June 1975).
3. C. Beckers, S. Chamberlain and P. Grimsrud, Quantitative methods for preliminary design of water quality surveillance systems. No. EPA-R5-72-001 (November 1972).
4. A. Cohen, Y. Bar-Shalom, W. Winkler and P. Grimsrud, A quantitative method for effluent compliance monitoring resource allocation, EPA contract No. 68-01-2232 (March 1975).
5. Areawide Waste Treatment Management Planning Agencies, 40 CFR Part 35, Federal Register Vol. 39, No. 93 (13 May 1974).
6. Water and Pollutant Source Monitoring, 40 CFR Part 35, Federal Register, Vol. 39, No. 68 (28 August 1974).
7. C. Beckers and S. Chamberlain, Design of cost-effective water quality surveillance systems, Report No. EPA-600/5-74-004 (January 1974).

(*Received September* 1975; *revised November* 1975)